WHAT IS CLAIMED IS:

1	1. A method for sound signal classification, comprising:
2	receiving a sound signal;
3	specifying meta-data to be extracted from the sound signal;
4	dividing the sound signal into a set of frames;
5	applying a fitness function to the frames to create a set of fitness data;
6	selecting a frame from the set of frames, if the frame's corresponding fitness
7	datum within the set of fitness data exceeds a predetermined threshold value;
8	extracting the meta-data from the selected frames; and
9	classifying the sound signal based on the meta-data extracted from the selected
10	frames.
1	2. The method of claim 1:
2	wherein the sound signal is a speech signal.
1	3. The method of claim 1 wherein specifying includes:
2	specifying age range meta-data.
1	4. The method of claim 1 wherein specifying includes:
2	specifying gender meta-data.
1	5. The method of claim 4 wherein selecting includes:
2	setting the threshold so that a ratio of frames selected to frames not selected is
3	between about 1:2 and about 1:3.

1	6. The method of claim 1 wherein specifying includes:
2	specifying accent meta-data.
1	7. The method of claim 1 wherein specifying includes:
2	specifying dialect meta-data.
1	8. The method of claim 1 wherein specifying includes:
2	specifying identity meta-data.
1	9. The method of claim 1 wherein dividing includes:
2	dividing the sound signal into a set of time frames.
1	10. The method of claim 1 wherein dividing includes:
2	dividing the sound signal into a set of equal length time frames.
1	11. The method of claim 1 wherein applying includes:
2	calculating a signal strength of the sound signal frame.
1	12. The method of claim 1 wherein selecting includes:
2	selecting a frame for meta-data extraction, if the frame's fitness datum exceeds
3	a greatest fitness datum within the set of fitness data by a predetermined margin.
1	13. The method of claim 1 wherein extracting includes:
2	extracting the meta-data from the selected frames using a Multi-Layer
3	Perceptron (MLP) neural network.

1 14. The method of claim 13 wherein extracting includes: 2 extracting the meta-data from the selected frames using a MLP neural network 3 having an input layer with nodes corresponding to the sound signal's Mel-Cepstral 4 components. 1 15. The method of claim 1 further wherein classifying includes: 2 assigning the sound signal to that meta-data class to which a largest number of 3 the selected frames have been assigned. 1 16. The method of claim 1 further wherein classifying includes: 2 adding together each of the selected frame's confidence scores for each meta-3 data class; and 4 assigning the sound signal to that meta-data class having a highest total 5 confidence score. 17. 1 The method of claim 1 further wherein classifying includes: 2 assigning the sound signal to that meta-data class having a statistically longest 3 run-length. 1 18. A method for sound signal classification, comprising: 2 receiving a speech signal; 3 specifying meta-data to be extracted from the sound signal; 4 dividing the sound signal into a set of equal length time frames; 5 applying a fitness function to the frames to create a set of fitness data;

6	selecting a frame for meta-data extraction, if the frame's fitness datum exceeds
7	a greatest fitness datum within the set of fitness data by a predetermined margin;
8	extracting the meta-data from the selected frames using a Multi-Layer
9	Perceptron (MLP) neural network;
10	adding together each of the selected frame's confidence scores for each meta-
11	data class; and
12	assigning the sound signal to that meta-data class having a highest total
13	confidence score.
1	19. A system for sound signal classification comprising a:
2	means for receiving a sound signal;
3	means for specifying meta-data to be extracted from the sound signal;
4	means for dividing the sound signal into a set of frames;
5	means for applying a fitness function to the frames to create a set of fitness
6	data;
7	means for selecting a frame from the set of frames, if the frame's
8	corresponding fitness datum within the set of fitness data exceeds a predetermined
9	threshold value;
10	means for extracting the meta-data from the selected frames; and
11	means for classifying the sound signal based on the meta-data extracted from
12	the selected frames.